Expertise & Skills (ES)

The CINECA «Find an expert» solution based on VIVO

Stefano Bolelli, Paola Galimberti, Loredana Rollandi
Core of our Research Portal : IRIS

- Institutional Research Information System (IRIS) is based on Dspace and initially developed by Cilea consortium (then included in CINECA consortium) in 2007 to manage publications and research outputs.

- IRIS is an extended and modular solution to progressively manage all research and related data, not only publications:
  - complete scientific curriculum of the researcher
  - research projects and activities
  - public engagements activities
  - scientific departments, research groups and center, laboratories, scientific equipment,...

- Aims: internal / external evaluation, monitoring and statistics, single point of collection of research data for different uses.
The need: public interface for all the data managed in IRIS

• Dspace has a public interface only for research outputs (publications)
• The need was the implementation of a public portal that:
  - aggregates and shows ALL the data managed in IRIS
  - allows to navigate and search transversely ALL type of data: people, units (e.g.: departments), projects and activities, expertise, outputs (publications and patents)
  - has a primary focus on the researcher:
    • “Find an expert” functionality
    • Researcher profile, that can be linked from other sites
The start of the project - Competence Repertoire (RC)

- The project for implementation of a Competence Repertoire (RC) starts in 2009 based on requirements of Vice-Rector and TTO.

- They needed a tool for researching skills in the University to satisfy requests from external stakeholders (journalists, companies, policy makers).

- The first Competence Repertoire for UNIMI was developed on a platform integrated with IRIS and released in 2010 after various stages of experimentation with stakeholders and revision cycles.

- It was not opened outside the University network.
Technology review: reasons for Vivo adoption

• Choice of the consortium for the public showcase of IRIS data:
  - based on a flexible data structure (RDF);
  - mapped on standard ontologies (as well as extensible for further needs);
  - open-source formats for display and interoperability of output data (Linked open-data)

• Extension of the data shown in the researcher's profile:
  - ALL curricular information and activities entered in IRIS are acquired and shown

• More flexible and integrable tool:
  - network of university research portals?
Technical activities

- Joint analysis between our University and Cineca for:
  - mapping between IRIS data schema and VIVO ontologies
  - reorganization of information displayed in VIVO interface (for each type of data we have defined a specific layout of page and redefined some navigation logic)
  - labels of each fields (for Italian and English interface)
  - search criteria and presentation of results

- Test on various intermediate prototypes

<table>
<thead>
<tr>
<th>IRIS LABEL</th>
<th>FIELD IN DATABASE (ODS SOURCE)</th>
<th>TARGET ENTITY</th>
<th>TARGET ENTITY TYPE</th>
<th>TARGET PROPERTY</th>
<th>LABEL UI IRIS-ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descrizione o denominatione del premio</td>
<td>ODS_L1_RM_PERSON.CV_PRIZE_DESCRIPTION</td>
<td>Prize</td>
<td><a href="http://vivoweb.org/ontology/core#Award">http://vivoweb.org/ontology/core#Award</a></td>
<td>rdfs:label</td>
<td>Description</td>
</tr>
<tr>
<td>Ente o istituzione che attribuisce il premio</td>
<td>ODS_L1_RM_PERSON.CV_PRIZE_DESCRIPTION</td>
<td>PrizeAwarder</td>
<td>foaf:Organization</td>
<td>rdfs:label</td>
<td>Confirmed by</td>
</tr>
<tr>
<td>Anno del premio</td>
<td>ODS_L1_RM_PERSON.CV_PRIZE_YEAR</td>
<td>PrizeAwardReceiptDate</td>
<td><a href="http://vivoweb.org/ontology/core#Date">http://vivoweb.org/ontology/core#Date</a></td>
<td>rdfs:label</td>
<td>Year Awarded</td>
</tr>
<tr>
<td>Motivazione del premio</td>
<td>ODS_L1_RM_PERSON.CV_PRIZE_MOTIVATION</td>
<td>PrizeAwardReceipt</td>
<td><a href="http://vivoweb.org/ontology/core#AwardReceipt">http://vivoweb.org/ontology/core#AwardReceipt</a></td>
<td>rdfs:label</td>
<td>Awards And Honors</td>
</tr>
<tr>
<td></td>
<td>ODS_L1_RM_PERSON.CV_FELLOWS.FELLOWSHIP_TYPE_DESC, INSTR(ODS_L1_RM_PERSON.CV_FELLOWS.FELLOWSHIP_TYPE_DESC,‘T’)</td>
<td></td>
<td></td>
<td></td>
<td>Tipo</td>
</tr>
</tbody>
</table>
From Vivo to UNIMI Expertise & Skills
Final steps of project

- Integration with the university portal (graphics elements and link to the university address book)

- Online user guide, privacy and legal notices, accessibility check

- Approved for production and open to the public at the end of December 2019

- Mail communication to internal list, press release
Expertise & Skills - Home Page

Different points of access to research skills and data

Free search

Lists of people / researchers
Lists of activities / projects
Lists of research areas
Lists of departments
Lists of results / research products
Expertise & Skills - Free search rules

- The free search is performed on the entity's "title" and on its significant properties (descriptive metadata like abstracts or keywords).

- The text entered by the user (e.g. climate change) is analyzed and the terms searched subsequently 1. as exact phrase, 2. with Boolean operator "AND", 3. with Boolean operator “OR” (progressive “relaxation”).

- To the results identified in the three steps will be given progressively decreasing importance (ranking). The other criteria used for "score" are: recurrence of searched terms and applied weight on some type of data.

- If the text is inserted between double quotes it is considered as a single phrase.

- The search engine extends the match also to terms that are derived from that root (stemming of terms). For example: to singular and plural terms, male and female, or other derivations.
Expertise & Skills - Free search results

The search is carried out in the descriptive cards of the various entities.

By clicking on the name, I access the complete profile.

**Matches**: allows a quick check of the total number and detail of the matches with the contents of the researcher's profile.
Areas and lines of research
Research projects
Publications and patents
Awards and scientific responsabilities
Other activities and Public Engagement
Cv (PDF and textual)
## Expertise & Skills: expertise classification

**LS - LIFE SCIENCES**

- **Concept Scheme**
  - Research Area 13 Associated Individuals

  - **LS1 - Molecular and Structural Biology and Biochemistry:** Molecular synthesis, modification and interaction, biochemistry, biophysics, structural biology, metabolism, signal transduction
    - **Concept**
      - Research Area 3 Associated Individuals
    - **LS1.1 - Molecular Interactions**
      - **Concept**
      - Research Area 38 Associated Individuals
    - **LS1.2 - General biochemistry and metabolism**
      - **Concept**
      - Research Area 35 Associated Individuals
    - **LS1.3 - DNA synthesis, modification, repair, recombination and degradation**
      - **Concept**
      - Research Area 6 Associated Individuals

### FILTER SEARCH

<table>
<thead>
<tr>
<th>ERC FIELD (2016)</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERC FIELD (2016)</td>
<td>312</td>
</tr>
<tr>
<td>ATECO</td>
<td>399</td>
</tr>
<tr>
<td>ISI - CRUI</td>
<td>111</td>
</tr>
</tbody>
</table>

### FILTER ERC FIELD (2016) BY ROOT ERC

- **PE - PHYSICAL SCIENCES AND ENGINEERING** 142
- **LF LIFE SCIENCES** 103
- **SH - SOCIAL SCIENCES AND HUMANITIES** 67
Expertise & Skills: some numbers

- Counting of managed entities:
  - 33 departments
  - 3,900 people
  - 2,700 research projects
  - 891 concepts
  - 225,000 publications

- Number of separate accesses from January to May 2020:

  ![Graph showing increase of searches on issues related to covid-19]

  Increase of searches on issues related to covid-19?
Expertise & Skills : main goals achieved

• Presentation:
  - A showcase of the skills and expertise of our researchers, where all researcher data managed in our IRIS are publicly navigable and searchable
  - The possibility for the single researcher to manage which data can be public and which has not to be visible online (for example: confidentiality agreements, or personal data in the CV).

• Search:
  - diversified weighting by type of outputs/publications
  - contextual and extensible linking between all the type of data
Expertise & Skills: future goals

- **SDG** integration for manual tagging of skills, in addition to ERC and local classification schemes
- Extension of the contents of researcher page to the **educational offer**:
  - VIVO entities have been introduced to represent: Course (e.g. degree course), Teaching (single), Professions (from Istat codes).
  - students as new stakeholders of the portal
- Integration of the **departmental sheet** with graphic informative summary data for all types of data
- **Collaboration networks/coauthors outside the University**

- Export of search results in a usable format (e.g.: csv)
- Export of usage statistics (and search strings, if possible) by the university admin

- Fingerprinting and automatic derivation of tags?
- Semantic search vs syntactic search?
Thanks!

Loredana Rollandi (loredana.rollandi@unimi.it)
IT administrator and analyst

Università degli Studi di Milano